- 75 -

SEOUENCE LISTING

(1)	GENERAL	INFORMATION

- APPLICANT: Nelson, Edward L. (i) Nelson, Peter J.
- TITLE OF INVENTION: NOVEL VECTOR FOR (ii) POLYNUCLEOTIDE VACCINES
- NUMBER OF SEQUENCES: 29 (iii)
- CORRESPONDENCE ADDRESS: (iv)
 - ADDRESSEE: MORGAN & FINNEGAN, L.L.P. (A)
 - STREET: 345 PARK AVENUE (B)
 - CITY: NEW YORK STATE: NEW YORK (C)
 - (D)
 - COUNTRY: USA (E)
 - ZIP: 10154 (F)
- COMPUTER READABLE FORM: (v)
 - MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB (A)
 - COMPUTER: IBM PC COMPATIBLE (B)
 - OPERATING SYSTEM: PC-DOS/MS-DOS (C)
 - SOFTWARE: WORDPERFECT 5.1 (D)
- CURRENT APPLICATION DATA: (vi)
 - APPLICATION NUMBER: PCT/US97/14306 (A)
 - FILING DATE: 14-AUG-1997 (B)
- PRIOR APPLICATION DATA: (vii)
 - APPLICATION NUMBER: US60/023931 (A)
 - FILING DATE: 14-AUG-1996 (B)
- ATTORNEY/AGENT INFORMATION: (viii)
 - NAME: KATHRYN M. BROWN (A)
 - REGISTRATION NUMBER: 34556 (B)
 - REFERENCE/DOCKET NUMBER: 2026-4236US1 (C)
- TELECOMMUNICATION INFORMATION: (ix)
 - TELEPHONE: (212) 758-4800 (A)
 - TELEFAX: (212) 751-6849 (B)
 - TELEX: 421792 (C)
- INFORMATION FOR SEQ ID NO:1: (2)
 - SEQUENCE CHARACTERISTICS: (i)
 - LENGTH: 453 base pairs
 - TYPE: nucleic acid (B)
 - STRANDEDNESS: single (C)
 - TOPOLOGY: unknown (D)
 - MOLECULE TYPE: cDNA (ii)

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No	
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:1:	
GGCCGCGTTG CTGGCGTTTT TCCATAGGCT CCGCCCCCT GACGAGCATC ACAAAAATCG ACGCTCAAGT CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC CTGGAAGCTCC CTCGTGCGCT CTCCTGTTCC GACCCTGCCG CTTACCGGAT ACCTCTCCGC CTTTCTCCCT TCGGGAAGCG TGGCGCTTTC TCAATGCTCA AGCTGGGCTG ATCTCAGTTC GGTGTAGGTC GTTCGCTCCA AGCTGGGCTG TGTGCACGAA CCCCCCGTTC AGCCCGACCG CTGCGCCTTA TCCGGTAACT ATCGTCTTGA GTCCAACCCG GTAAGACACG ACTTATCGCC ACTGGCAGCA GCCACTGGTA ACAGGATTAG ACTTATCGCC ACTGCGCTA CAC ACTACGGCTA CAC ACGACCCCCTTA TCCGCCCTA ACTTATCGCC ACTACGGCTA ACAGGATTAG CAGAGCGAGG TGCTACAGA CTTCTTGAAG TGGTGGCCTA ACTACGGCTA CAC	40 80 120 160 200 240 280 320 360 400 440 453
(2) INFORMATION FOR SEQ ID NO:2:	
(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 453 base pairs(B) TYPE: nucleic acid	
(C) STRANDEDNESS: single (D) TOPOLOGY: unknown	
(ii) MOLECULE TYPE: cDNA	
(iii) HYPOTHETICAL: No	
(iv) ANTI-SENSE: No	
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:2:	
GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG GCTGAACGGG GGGTTCGTGC ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA CCGTGAGCAT TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGGCGGACA	40 80 120 160 200 240 280
ACCCCCACC CTCGCACA GAGAGCGCAC	320

360

400

440 453

- INFORMATION FOR SEQ ID NO:3: (2)
 - SEQUENCE CHARACTERISTICS: (i) (A) LENGTH: 209 base pairs

GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCGCAC

GAGGGAGCTT CCAGGGGGAA ACGCCTGGTA TCTTTATAGT

CCTGTCGGGT TTCGCCACCT CTGACTTGAG CGTCGATTTT

TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC

CAGCAACGCG GCC

	(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: unknown	
(ii)	MOLECULE TYPE: cDNA	
(iii)	HYPOTHETICAL: No	
(iv)	ANTI-SENSE: No	
(xi)	SEQUENCE DESCRIPTION:SEQ ID NO:3:	
GGATTCGAAC CTGCTCCCTT	GGACTTTTGA AAGTGATGGT GGTGGCCGAA CTTCGAAGTC GATGACGGCA GATTTAGAGT TGGCCGCTCG GGAACCCCAC CACGGGTAAT GCCTGCTCCC TTATCGGGAA GCGGGCGCA TGACGCGCCG CTGTAAAGTG TTACGTTGAG	40 80 120 160 200 209
(2) INFORM	MATION FOR SEQ ID NO:4:	
(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 209 base pairs	
	(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: unknown	
(ii)	MOLECULE TYPE: CDNA	
(iii)	HYPOTHETICAL: NO	
(iv)	ANTI-SENSE: No	
(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 4:	
TTGATATGAT AGTAAAAGCA AGGGAGCAGA	TCAACGTAAC ACTTTACAGC GGCGCGTCAT GCGCCCCGCT TCCCGATAAG GGAGCAGGCC TTACCCGTGG TGGGGTTCCC GAGCGGCCAA CTCTAAATCT GCCGTCATCG ACTTCGAAGG TCCCCCACCA CCATCACTTT CAAAAGTCCG	40 80 120 160 200 209

INFORMATION FOR SEQ ID NO:5: (2)

- SEQUENCE CHARACTERISTICS: (i) LENGTH: 6 base pairs
 TYPE: nucleic acid (A)
 - (B)
 - STRANDEDNESS: single (C)
 - TOPOLOGY: unknown (D)
- MOLECULE TYPE: cDNA (ii)

	(iii) HYPOTHETICAL: No	
	(xi) SEQUENCE DESCRIPTION:SEQ ID NO:5:	
AATA	AAA	6
(2)	INFORMATION FOR SEQ ID NO:6:	
	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 6 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown 	
	(ii) MOLECULE TYPE: cDNA	
	(iii) HYPOTHETICAL: No	
	(iv) ANTI-SENSE: No	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:	
ATT	AAA	6
(2)	INFORMATION FOR SEQ ID NO:7:	
	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 6 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown 	
	(ii) MOLECULE TYPE: cDNA	
	(iii) HYPOTHETICAL: No	
	(xi) SEQUENCE DESCRIPTION:SEQ ID NO:7:	
AGT.	'AAA	6
(2)	INFORMATION FOR SEQ ID NO:8:	
(2)	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 6 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown	
	(ii) MOLECULE TYPE: cDNA	

	(iii) HYPOTHETICAL: No	
	(xi) SEQUENCE DESCRIPTION:SEQ ID NO:8:	
AAGA	AC	6
(2)	INFORMATION FOR SEQ ID NO:9:	
	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 6 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown 	
	(ii) MOLECULE TYPE: cDNA	
	(iii) HYPOTHETICAL: No	
	(iv) ANTI-SENSE: No	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:	
AATA	CA	6
(2)	INFORMATION FOR SEQ ID NO:10:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 227 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown	
	(ii) MOLECULE TYPE: cDNA	
	(iii) HYPOTHETICAL: No	
	(iv) ANTI-SENSE: No	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:	
GCGG ACAG AGTG ACCA	GCCTTG TCCTAATAAA ATTAAGTTGC ATCATTTTGT 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	40 80 20 60 200 27
(2)	INFORMATION FOR SEQ ID NO:11:	
	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 227 base pairs(B) TYPE: nucleic acid	

(C) STRANDEDNESS: single (D) TOPOLOGY: unknown	
(ii) MOLECULE TYPE: cDNA	
(iii) HYPOTHETICAL: No	
(iv) ANTI-SENSE: No	
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:11:	
ATAATATTAT AGAGGACACC TAGTCAGAAC AAATGATGCA ACTTAATTTT ATTAGGACAA GGCTGGTGGG CACTGAGTG GCAACTTCCA GGGCCAGGAG AGGCACTCAC TACTCATCCT GCAGGGGAAG GACGGGCATT GGCCTGTGCT GCCCGGGGGC TCTGACTACA GGTCTCCCC ATCCCCGCCT GGGGTCAAGG CATCCACTCA CCATATGGCC CTTAAGG	40 80 120 160 200 227
(2) INFORMATION FOR SEQ ID NO:12:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 252 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown	
(ii) MOLECULE TYPE: cDNA	
(iii) HYPOTHETICAL: No	
(iv) ANTI-SENSE: No	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:	
CCTCGGTACC TGCCATGGCG CGGATTCTTT ATCACTGATA AGTTGGTGGA CATATTATGT TTATCAGTGA TAAAGTGTCA AGCATGACAA AGTTGCAGCC GAATACAGTG ATCCGTGCCG GCCCTGGACT GTTGAACGAG GTCGGCGTAG ACGGTCTGAC GACACGCAAA CTGGCGGAAC GGTTGGGGGT GCAGCAGCCG GCGCTTTACT GGCACTTCAG GAACAAGCGG GCGCCTTAAG GGCCATATGC CG	40 80 120 160 200 240 252
(2) INFORMATION FOR SEQ ID NO:13:	

- SEQUENCE CHARACTERISTICS: (i) LENGTH: 35 base pairs TYPE: nucleic acid (A) (B)
 - (C) (D) STRANDEDNESS: single TOPOLOGY: unknown

	(ii)	MOLECULE TYPE: cDNA	
	(iii)	HYPOTHETICAL: No	
	(iv)	ANTI-SENSE: No	
(xi)	SEQUE	NCE DESCRIPTION:SEQ ID NO:13:	
CCTCC	GGTACC T	TGCCACCATG GCGCGGATTC TTTAT	35
(2)	INFORM	ATION FOR SEQ ID NO:14:	
	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 38 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown	
	(ii)	MOLECULE TYPE: cDNA	
	(iii)	HYPOTHETICAL: No	
	(iv)	ANTI-SENSE: No	
	(xi)	SEQUENCE DESCRIPTION:SEQ ID NO:14:	
CGGC	ATATGG	CCTTAAGGCG CCCGCTTGTT CCTGAAGT	38
(2)	INFORM	NATION FOR SEQ ID NO:15:	
	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 228 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown	
	(ii)	MOLECULE TYPE: cDNA	
	(iii)	HYPOTHETICAL: No	
	(iv)	ANTI-SENSE: NO	
	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:15:	
GCCI	TTAAGGG GGATGG	CCATATGGTG AGTGGATGCC TTGACCCCAG GGGAGACCTG TAGTCAGAGC CCCCGGGCAG	4 (8 (
GAGT CACC	TGCCTCT CAGCCTT	TGCCCGTCCT TCCCCTGCAG GATGAGTAGT CCTGGCCCTG GAAGTTGCCA CTCCAGTGCC GTCCTAATAA AATTAAGTTG CATCATTTTG TGTCCTCTAT AATATTAT	120 160 200 228

INFORMATION FOR SEQ ID NO:16: (2)

- (A) LENGTH: 1425 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown
- MOLECULE TYPE: cDNA (ii)
- HYPOTHETICAL: No (iii)
- ANTI-SENSE: No (iv)
- SEQUENCE DESCRIPTION: SEQ ID NO:16: (xi)

TGCCATGGCG CGGATTCTTT ATCACTGATA AGTTGGTGA 40 CATATTATGGC TATACAGTGA TATACAGTGA 80 AGTTGCAGC GAATACAGTG ATCCGTGCCG GCCCTGGACT 120 GTTGACGACA GTTGGCGAAC GCTGGCGTAG ACGGTCTGAC GCCCTGACA 160 CTGGCGAAC GGTTGGGGGT GCACCACCGC GCCCTTTAGC 240 GGCACTTCAG GACCAGGCGT GGCCATTTGG 240 TGAGTGGATG CCTTGACCCC AGGCGGGGAT GGCCATATGG 240 TGAGTGCAGA GCCCCCGGGC AGCACAGCCC AATGCCCCTC 320 CTTCCCCACT CCCCAGTGCC TCTCCTGGCC CTGGGAAGTCC 320 TGTGACCCT CCCCACCAGC CTTGTCCTAA AATGCCCCTC 360 TGTGACCCT CCCCACCAGC CTTGTCCTAA AATATTAAG 40 TAGGCGCGGT ATTGTCAAAA AGTTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA						
CATATATATE TARACTEA ATCCGTGCC GCCCTGGACT 120 GTTGAACGAG GTCGGCGTAG ACGGTCTGAC GCCCTGGACT 120 GTGGCGAAC GGTTGGGGGT GCACCTCTAC GAACAGCGAAA 160 GGCACTTCAG GAACAGCGG GCGCCTTAAG 200 GGACATCAGAG GCCCCTGGCC AGGCGGGAT 240 TGTAGTCAGA GCCCCCCGGC AGACAGGCC 280 ATTCCCCTC AGTGGATAGT GCCCCCCGTC 320 TGTAGTCAGA GCCCACCAGC CTTGCCCGG GTGGGATACC 360 TGTGACCCT CCCCAGTGC CTTCCTCGC CTGGAAGTTGC 360 TGTGACCTATT TTGTCTGACT AGTTCCCCGG GTGGATTCC 360 TTGCATCATT TTGTCTGACT TCTCCTCAA AACACTTTAA 440 TCGATCATAT ATTGATATTATATATATATA 480 AACACTTTAA 480 TCGACTCGA CAATACCCCA AGGTTCACCA GCTTCCCCA CCTGTCACA 640 TCGACTTCA AGACTCTAAA TCTGCCCCA CCTGTACCAC 680	TGCCATGGCG	CGGATTCTTT	ATCACTGATA	AGTTGGTGGA		
AGTTGCAGCC GAATACAGTG ATCCGTGCCG GCCTGGACT 120 GTTGAACGAG GTCGCGTTAG ACGTCTGAC GACACGCAAA 160 CTGGCGAAC GGTTGGGGGT GCAGCAGCCA GCCCTTACT 200 GGCACTTCAG GAACAAGCGG GCCCCTTAAG GCGCTTTACT 200 TGAGTGAGTAG CCTTGACCCC AGGCGGGGAT GGGGGAGACC 280 TGTAGTCAGA GCCCCCGGCC AGGCGGGGAT CCTGCAGCC 320 CTTCCCCTGC AGTGAGTAGT GACTGCCCGG GTGGATCCC 360 CTTGACCCT CCCCAGTGCC CTTGCTCAA AATGCCCGTC 400 CCACTCCAGT CCCCACACCC CTTGCTCAA AAATTATAAG 40 TGAGTTGAT ATCGAATTCT TTCTCACAGT TATAATTATAA 480 AAGGGGCAG CCAATTAATA GACTTTAAA AACACTTTAA 520 AGCGGCGCT CAAAGGGAGC AGACTCTAAA TCTGCCGCA TGGTGGGGT 560 CCGAGCGCGC CAAAGGAGC AGACTTCAC TGGTGGGGT 60 60 CCGAGCGGC	CATATTATGT	TTATCAGTGA	TAAAGTGTCA	AGCATGACAA		
GTTGGACGAG GTTGGACGA ACGTTGAC 200 CTGGCGGAAC GGTTGGGGGT CAGCAGCCG GCGCTTTAGT 240 GGCACTTCAG GAACAAGCGG GCGCCTTAAG GCGCATATGG 240 TGAGTGGATG CCTTGACCC AGGCGGGGT GGGCATATGG 240 TGTAGTCAGA GCCCCAGCC AGGCACAGGC AATGCCCGTC 320 CTTCCCCTGC AGTGAGTAGT GACTGCCCGG GTGGAATTCC 360 TGTGACCCT GCCCACAGC CTTGTCCTAA AATGCCCGTC 360 TGTGACTATT TTTGTCTGACT AGGTGTCCCC CTGGAAGTTG 400 TGAGACTATT TTTCTCTGACT AGGTTTCCCC CTTGTCCTAA 440 TGAGCGCGCT CATTTGAAT TTCTCAACGT AACACTTTAC 520 AGGGAGCAG CAAAGGGAGC GCCATTAACCG GCTTCCCCAT 560 CCCGAGCGCC CAAAGGGAGC AGACTCTAAA TCTGCCGCAT 640 CCCGAGCGCC CACAAGAAA TCTCTCCCCA CCACCACCCC CCACCACCCC GCTGGAGCAT 720 TAATTCAAAAA <td< td=""><td>AGTTGCAGCC</td><td>GAATACAGTG</td><td>ATCCGTGCCG</td><td>GCCCTGGACT</td><td></td><td></td></td<>	AGTTGCAGCC	GAATACAGTG	ATCCGTGCCG	GCCCTGGACT		
CTGGCGGARC GATTGGGGGT GCACCTTCAG GAACAAGCGG GCGCCTTAAG GGCCATTATG 240 TGAGTGGATG CCTTGACCCC AGGCGGGGAT GGCCATTATG 280 AGGCGGGAT 320 TGTAGTCAGA GCCCCCGGGC AGCACAGGCC CATGCCCGG GTGGAGTCCC 360 TGTGACCCT CCCCAGTGCC TCTCCTGGCC CTGGAGATTG 400 CCACTCCAGT GCCCACCAGC CTTGTCCTAA 440 CCACTCCAGT ACCACTTCAT TAAAATTAAG 440 CCAGTGGCT CATTTGATT AGGTGCCCC GCTTGCCCAT TAAAATTAAG 440 AGGGGCGGC CATTTGATAT GATGCGCCCC GCTTCCCCAT TAAAATTAAG 440 ACGCGGCGCT CATTTGATAT GATGCGCCCC GCTTCCCGAT ACACCTTTAC 520 AAGGGAGCAG CATTTGATAA GCTTCCCCAA ACACCTTAA 640 640 CCGAGCGGC CAAAGGGAGC CACACCTCAA GCACCCCTAA 640 640 TTTCAAAAGT CCTGCACCACCC CCTGCTCACCACCCCACCACCACCACCACCACCACCACCACCAC	GTTGAACGAG					
GGCACTTCAG GAACAGGGG GGCCTTAAG GGCGACTTCAG 280 TGAGTGGATG CCTTGACCCC AGGGGGGGAT 320 CTTCCCCTGC AGTGAGTAGT GACTGCCGG GTGGATCCC 360 CTTGCCCAGT CCCCCAGTGC CTCCTGGCC CTGGAAGTTG 400 CCACTCCAGT GCCCACCAGC CTTGTCCTAA TAAAATTAAG 440 TTGCATCATT TTGTCTGACT AGGTGTCCCC GCTGGAAGTTG 520 AGGGGGCGC CATTTGATAT TACGAATTACC 520 AGGGGAGCAG CCATTGAACA ACACTTTAC 520 AAGGGAGCAG CCAAAGGAGC CGTTCCCGAT 600 CCGGAGCGC CAAAGGGAGC AGACTCTAAA TCTGCCGAT 600 CCGAAGCAGC CCAAAGGAGC CCACCATCAC 680 TTTCAAAAGT CCGGAAGAAT TCTGCCACC CCACCATCAC 680 TTTCAAAAGT CCTGCTAACC CGTTTACCGG 720 CATACCTCGC TCTGCTAATC CTGTTACCGG 720 CAGTGGCGAT ACACCACTCCC CTGCTAAGAA 840	CTGGCGGAAC	GGTTGGGGGT				
TGAGTGGATG CCTTGACCCC AGCACAGGCC 320 TGTAGTCAGA GCCCCCGGGC AGCACAGGCC 360 CTTCCCCTGC CCCCAGTGCC TCTCCTGGCC CTGGAGTTG 400 TGTAGCCCT CCCCAGTGCC TCTCCTGGCC CTGGAGTTG 400 TGCATCATT TTGTCTGACT AGGTGTCCTC TAAAATTAAG 440 TTGCATCATT TTGTCTGACT AGGTGTCCCC TTTAAAATTAA 480 TAGGGGGCGCT CATTTGATAT GATGGCCCC GCTTCCCGAT 520 AGGGGGGCGC CATTTGATAT GATTACCCG TGGTGGGGTT 560 AAGGGAGCAG CCCAGTAAAA GCATTACCCG TGGTGGGGTT 640 TCGACTTCGA AGACTCTAAA TCTGCCGTCA 640 TCGACTTCGA AGACTCTTAA TCTGCCGTCA 640 TTTCAAAAGT CCCACACATCAC CGTTAGCCG 720 TAATCTGC TCTGCTAATC CTGTTACCAG 720 CATTCCTGC TCTGCTAATC TTACCGGGTTG GCACCGCTA 760 CACCACATCGC TCTGCAGACCC GCGCTGAACAG	GGCACTTCAG	GAACAAGCGG	GCGCCTTAAG			
TGTAGTCAGA GCCCCGGG GACGAGGCC GTGGGATCCC 360 TGTGACCCT CCCCAGTGCC TCTCCTGGCC CTGGAAGTTG 400 CACTCCAGT GCCCACCAGC CTTGTCCTAA TAAAATTAAG 440 TTGCATCATT TTGTCTGACT AGGTGTCCTC TATAATATTA 480 TAAGCTTGAT ATCGAATTCT TCTCACACGT AACACCTTTAC 520 AGGGGCGCGC CATTGGAAT GATCCCCC GCTTCCCGAT 560 AGGGAGCAG CAAAGGGAGC AGACTCTAAA TCTGCCGTCA 640 TCGACTTCGA AGGTTCGAAT CTTCCCCCA CCACCATCAC 680 TCTCAAAAGT CCGAAAGAAT TCCTGCAGCC CGTGTAGCCG 720 TAGTTAGGCC CACACTCAA GAACTCTGAA 760 CATACCTCGC TCTGCTAATC CTGTTACCAG TGGCTGCC 800 CATACCTCGC TCTGCTAATC CTGTTACCAG TGGCTGCC 800 CATACCTCGC TCTGCTAATC CTGTTACCAG TGGCTGCC 800 CACTGGCGAT AAGTCGTGTC TTACCAG TGGCTGCC 880 CGGGTTCGTG CACACAGCC AGCTTGAAC GAACTCACAC 880 CGATAGTTAC CGGATAAGAC GCACGGCTA 760 CGCACGACTC CACACAGCC AGCTTGAAC 760 CACCGAACTG CACACAGCC AGCTTGAAC 760 CGCACGCTTC CCGAAAGGAA 760 TAAGCGCAC GCGCTAC AGCTTGAAC 760 CCCGAAGGAA AAAGCCC AGCTTGAAC 760 CCCGAAGGAA AAAGCCC AGCTTGAAC 760 CCCGAAGGAA 760 CCCGAAGGAA 760 CCGGATAAGGC GCAGGGTCG GGCTGAACGG 880 CGCCACGCTTC CCGAAAGAAC 760 CCCGAAGGAA AAAGCCCC AGCTTGAAC 760 CCCGAAGGAA 760 CCCGAAGGAAC 760 CCCCTCACGAACGC 760 CCCGAAGGAAC 760 CCCCTCAGAGAAC 760 CCCCTCAGAGAAC 760 CCCCTCAGAGAAC 760 CCCCTCAGAGAAC 760 CCCCTCAGAGAAC 760 CCCCTCAGAGGAAC 760 CCCCTCAGAGAAC 760 CCC	TGAGTGGATG	CCTTGACCCC				
TGTGACCCCT CCCCAGTGCC TCTCCTGGCC CTGGAAGTTG 400 CCACTCCAGT GCCCACCAGC CTTGTCCTAA TAAAATTAAG 440 TTGCATCATT TTGTCTGACT AGGTGCCCCC TATAAAATTAAG 440 AGCGGCGCGT CATTTGATAT TCTCCAACGT AACACTTTAC 520 AGGGAGCAG GCCAGTAAAA GCATTACCCG GCTTCCCGAT 600 CCCGAGCGGC CAAAGGGAGC AGACTCTAAA TCTGCATCAA 640 CCCGAGCGGC CAAAGGAAC CCTTCCCCCA CCACCATCAC 680 TTGCATTCGA AGGTTCGAAT TCCTGCAGCC CGTTCCCGAT 680 TTTCAAAAGT CCGAAAGAAT TCCTGCAGCC CGTTAGCCG 720 TAGTTAGGCC CCCACTCAA GAACTCTTAA GCACCGCCTA 760 CATACCTCGC TCTGCTAATC CTGTTACCAG TGGCTGCCTA 760 CAGTGGCGAT AAGTCGTGTC TTACCGGGTT GGACTGCC AGCTCAAGA 840 CGATAGTTAC CGGATAAGGC GCAGCGGTCG GGCTGAACGG 880 CAGTGGCGAT AAGTCGTGC AGCTTGGAGC GAACGACCTA 920 CACCGACTC CCCAACAGCC AGCTTGAAC GAACCACCGCCTA 920 CACCGACTC CCCAACGCCC AGCTTGGAGC GAACGACCTA 920 CACCGACTC CCCGAAGGGAG AAAGGCGGAC AGGTATCCGG 1000 TAAGCGGCAG GGTCGGAACA GGAGAGCCA AGGTATCCGG 1000 TCCAGGGGGA AACGCCTGT ATCTTTATAG TCCTGTCGGG 1080 TCCAGGGGGA AACGCCTA TCGAAAAACG CGAGGAACGC 1040 CCCGAAGGGAG AACGCCTA TCGAAAAACG CAGCCCC 1160 CGCCGGGGGA TCCCGAAGGAC TCCCTCTAAC TGAAAAACG CAGCCCC 1160 CGCCGGGGGA TCCCGGAGCC TCCGATTT TTGTGATGCT 1120 CGTCAGGGGG GCCGAGCCTA TGGAAAACC CCAGCAACCC 1160 CGCCGGGGGA TCCCGTTTTTCC CCACCCCAACCC 1160 CGCCGGGGGA TCCCGTTTTTCC CCACCCTAAGA 1200 CGCCAGCTCT TCGCTTTTCT TCCCTCTAAC TGAAAAACC CAGCCAACCC 1160 CGCCGGGGGA TCCCGTTTTTCC CCACCTCTAACCC 1240 AGTTTTCTTT TCCCTTTTTT TCCGTTTTCC TTTTTCCCCCC 1240 CGCCAATGCT TGGTTGCTAT TTTTGGAAACC CCCCTTAAGGC 1320 CGCCAATGCT TGGTTGCTAT TTTTGGAAACT CCCCTTAAGGC 1320 CGCCAATGCT TGGTTGCTAT TTTTGGAAACT CCCCTTAAGGC 1320 CGCCAATGCT TGGTTGCTAT TTTTGGAAACT CCCCTTAAGGC 1360 CGCCAATGCT TGGTTGCTAT TTTTGGAAACT CCCCTTAAGGC 1320 CGCCAATGCT TGGTTGCTAT TTTTGGAAACT CCCCTTAAGGC 1360 CGCCAATGCT TGGTTGCTAT TTTTGGAACCT CCCCTTAAGGC 1360 CGCAAGGGATT CCCTGCAGAGC ACCCTTAAGACC 14400	TGTAGTCAGA	GCCCCGGGC				
TOTACCCCT CCACAGC CCACTCCAGT CCACTCCAGT CCACTCCAGT TTGCTCACT TTGCTCACT TTGCTCACT TTGCTCACT TTGCTCACT TTGCTCACT TTGCTCACT ACCGACCGC CCATTCATT TTCCCAACGT ACCGCCCCC CCATTCAAAA ACCCTTTAC ACCGACCGC CCAAGGGAC CCAAGGGAC CCAAGGGAC CCAAGGGAC CCCAAGGGAC CCCAAGGGAC CCCAAGGGAC TCGACTTCAA AGGTTCCAAA AGGTTCCAAA ACACTTTAC CCGACCGC CCAAAGGAC TCCGACTCCA AGGTTCGAAT TCCTCACCC CCGACCGC CCAAAGGAC TCCCCCCC CCCACCCTCA ACCACTTCAA ACCACTTCAA ACCACTTCAA ACCACTTCAA ACCACTTCAA ACCACTTCAA ACCACTTCAA ACCACTTCAA ACCACTTCAA CCACTCCCC CCTTCCCCCA CCACCACCC CCGTTAGCCC CCGTTAGCCC CCGTTACCCCC CCGTCAACGC CCGTTACCCCC CCGTTACCCCC CCGTCAACCC CCGTTACCCCC CCGTCAACCC CCGTTACCCCC CCGTCAACCC CCGTTACCCCC CCGTCAACCC CCGTTACCCCC CCGTCACCCC CCGCTCACCC CCGCTCA CCGACCCCCC CCGCTCACCC CCGCACCCCC CCCACCCCCC CCCACCCCCC CCCACCCCCC	CTTCCCCTGC	AGTGAGTAGT	GACTGCCCGG			
TTGCATCATT TTGTCTGACT AGGTGTCCTC TATAAATATA 480 TAAGCTTGAT ATCGAATTCT TTCTCAACGT AACACTTTAC 520 AGCGGCGCGT CATTTGATAT GATGCCCCCCCCCACCCCCACCCCCACCCCCCACCCCCCCC	TGTGACCCCT	CCCCAGTGCC	TCTCCTGGCC	-		
TTGCATCATT TTGTCTGACT TAGCTTGAT ATCGAATTCT TTCTCAACGT AGCGGCGCGT CATTTGATAT GATGCGCCCC GCTTCCCGAT AGCGGCGCGC CCAGCGGC CCAGAGGGC CCAGAGGGC CCAAAGGGAGC CCCAGCGGC CCAAAGGGAGC CCGAACGAT TTTCAAAAGT TTTCAAAAGT TTCTCAAAAGT TTCTCAAAAGT TTCTCAAAAGT TTCTCAAAAGT TTCCACAAAGT TCTCCCCA CCGAACGACC CCGAAAGAAT TCCTCCCCA CCACCATCAC CACCATCAC GAACTCTGA GAACTCTGAA GAACTCTGTA GCACCCCCA CCGTTAGCCG CCGAACGAC CCGAAAGAAT TCCTCACCC CCGTTAGCCG CCGTTAGCCG CCGTTAGCCG CCGACGCCT TTACCAGC CCGTTACCAG CCGCCTA TGCTCACC CCGTTACCAG CCGCCTA TGCCCCCA CCGCCTA TGCCCCCA CCGCCCTA TGCCCCCA CCGCCCTA TGCCCCCA CCGCCCTA TGCCCCCC CCGCCCTA TGCCCCCC CCGCCCTA TGCCCCCC CCGCCCTA TGCCCCCC CCGCCCTA TGCCCCCC CCGCCCCC CCGCCCTA TGCCCCCC CCGCCCCC CCCCCC	CCACTCCAGT	GCCCACCAGC				
TAAGCTTGAT ATCGATTCT AGGGGCGCGT CATTTGATAT GATGCGCCCC GCTTCCCGAT AGGGAGCAG GCCAGTAAAA GCATTACCG TGGTGGGGTT CCCGAGCGGC CAAAGGAGC AGACTCTAAA TCTGCCGTCA CCGAGCGGC CAAAGGAGC AGACTCTAAA TCTGCCGTCA CTGACTTCGA AGGTTCGAAT CCTTCCCCCA CCACCATCAC CAGTGCGCA ACCACTTCAA GACTCTGTA GCACCGCTA CATACCTCGC TCTGCTAATC CTGTTACCAG GCACCGCCTA CAGTGGCGAT AAGTCGTGCT TTACCAGGGTT GGACTGCTGC CAGTGGCGAT AAGTCGTGCT TTACCAGGGTT GGACTCAAGA CGATAGTTAC CGGATAAGGC GCAGCGGTCG GGCTGAACGG CACCGAACTG CACACAGCCC AGCTTGAGC GAACCGCCTA CACCGAACTG AGATACCTAC AGCGTGAGCA TTGAGAAAGC CACCGAACTG AGATACCTAC AGCGTGAGCA AGGTATCCGG CCACGGCTC CCGAAGGAG AAAGGCGGAC AGGTATCCGG TAAGCGGCAG GGTCGGAACA GGAGAGCGC AGGTACCGG TCCAGGGGGA AACGCCTGGT ATCTTATAG TCCTGTCGGG TTCCAGCGGGA AACGCCTGT TTTTTATAG TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGATT TTTGTGATGCT CCTCAGGGGGA TCCGGAGCC TCGACATACC CCAGCAACGC GGCCGGGGAA TCCCGAGAGCC TCGACTTAGAAAACG CCAGCAACGC GGCCGGGGAACT TCGGAAAAACG CCAGCAACGC GGCCGGGGGA TCCCGAGAGC TCACTCTAGA TGAGAAAACG CGCCGGGGGA TCCCGAGAGC TCCTTATAG TCCTGTCCGG TCTGACTTGA GCGTCGATT TTTGTGATGCT TTTGCCACC TCTGACTTGA GCGTCGATT TTTGTGATGCT TCGGCAGGGG GCGGAGCCTA TGGAAAAACG CCAGCAACGC GGCCGGGGAA TCCCGAGAGCC TCACTCTAGA TGAGAAAACG CCAGCAACGC TCGACTTCA TCGGAAAAACG CCAGCAACGC TCAGCGGGGA TCCCGAGAGCC TCACTCTAGA TGAGAGAGCA TCCCGAACGC TCAGCTTTCTTT TCCGTTTTGT GCAATTTCC TTATGATACC TTATGATACC GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG GAGCCAATGCT TGGTTGCTAT TTTGGAACCT TTATGATACC GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG GAGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG GAGCCAATGCC AACTGCC AACTGC	TTGCATCATT	TTGTCTGACT	AGGTGTCCTC			
AAGGGGGCGCT AAAGGAACA CCCGAGCGCC CCAAAGGAACA TCCGAGCGCC CCAAAGGGAGC CCCAAAGGAACA TCCGCCCCA TCCGACTTCCA AGGTTCGAAT TCCTCCCCCA TCTCCAAAAGT TCCTCCCCA CCACCATCAC CCGAAAGAAT TCCTGCAGCC CATACCTCCC CATACCTCCC CATACCTCCC CATACCTCCC CATACCTCCC CACCACTCAA CCACTTCAA CCACTTCAA CCGATACCTCC CACCACTCAC CACCACCCC CACCACCATCAC CCGAAAGAAT TCCTGCAGCC CACCACCCC CACCACCACCAC CCACCACCAC CCACCA	TAAGCTTGAT	ATCGAATTCT	TTCTCAACGT			
AGGGAGCAGC CCCAAAGGAACC CCCAAAGGAACC AGACTCTCAA AGGTTCGAAT TTCAAAAAGT CCCGAAAGAAT TCCTGCCCCA CCACCATCAC TTGCACTTCCA AGGTTCGAAT TCCTGCAGCC CCACCATCAC TTGCTAAGCC CCACCATCAC CCACCATCAC TCGCTACCC CCACCATCAC TCGCTCCCCA CCACCATCAC TCGCTAACC CCACCATCAC CCACCATCAC TCGCTACCC CCACCATCAC TCGCTACCC CCACCATCAC CCACCATCAC TCGCTACCC CCACCATCAC CCACCACCAC CCACCATCAC CCACCACCCC CCACCACCAC CCCACCACCC CCACCA	AGCGGCGCGT	CATTTGATAT				
TCGACTTCGA AGGTTCGAAT CCTTCCCCCA CCACCATCAC TTTCAAAAGT CCGAAAGAAT TCCTGCAGCC CGTGTAGCCG TAGTTAGGCC ACCACTTCAA GCACTCTGTA GCACCGCCTA TAGTTAGGCC ACCACTTCAA GCACTCTGTA GCACCGCCTA TAGTTAGGCC TCTGCTAATC CTGTTACCAG TGGCTGCTGC CAGTGGCGAT AAGTCGTGTC TTACCGGGTT GGACTCAAGA 840 CGATAGTTAC CGGATAAGGC GCAGCGGTCG GGCTGAACGG GGGGTTCGTG CACACAGCCC AGCTTGGAGC GAACGACCTA 920 CACCGAACTG AGATACCTAC AGCGTGAGCA AGGTATCCGG GCACGCTTC CCGAAGGAG AAAGGCGGAC AGGTATCCGG 1000 TAAGCGGCAG GGTCGGAACA GGAGAGCGA CGAGGGAGCT 1040 TCCAGGGGGA AACGCCTGGT ATCTTTATAG TCCTGTCGGG 1080 TTTCGCCACC TCTGACTTGA GCGTCGATT TCGTGATGCT 1120 CGTCAGGGGG GCGGAGCCTA TGGAAAACG CCAGCAACGC 1160 GGCCGGGGGA TCCGGAGAGC TCACTCTAGA TGAGAGAGCA 1200 GTAGGGGAG GACAGAGACT CGAATTTCCC GAGCTATTTC 1240 AGTTTCTTT TCCGTTTTGT GCAATTTCCC TTATGATACC 1280 GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG 1320 GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG 1320 GAGGCCCTC AACTGGCCT ATAAAGGGCC AGCCTGACCT 1360 GCCAATGCT TGGTTGCTAT TTTTGGAAACT CCCCTTAGGG 1320 GAGGCCAATGCT TGGTTGCTAT TTTTGGAAACT CCCCTTAGGG 1320 GAGGCCCTC AACTGGCCT ATAAAGGGCC AGCCTGACCT 1360 GCCAATGCT TGGTTGCTAT TTTTGGAAACT CCCCTTAGGG 1320 GAGGCCCTC AACTGGCCT ATAAAGGGCC AGCCTGACCT 1360	AAGGGAGCAG	GCCAGTAAAA				
TTTCAAAAGT CCGAAAGAAT TCCTGCAGCC CGTGTAGCCG 720 TAGTTAGGCC ACCACTTCAA GAACTCTGTA GCACCGCCTA 760 CATACCTCGC TCTGCTAATC CTGTTACCAG TGGCTGCTGC 800 CAGTGGCGAT AAGTCGTGTC TTACCGGGTT GGACTCAAGA 840 CGATAGTTAC CGGATAAGGC GCAGCGGTCG GGCTGAACGG 880 GGGGTTCGTG CACACAGCCC AGCTTGGAGC GAACGACCTA 920 CACCGAACTG AGATACCTAC AGCGTGAGCA TTGAGAAAGC 960 GCCACGCTTC CCGAAGGGAG AAAGGCGGAC AGGTATCCGG 1000 TAAGCGGCAG GGTCGGAACA GGAGAGCGC CGAGGGACT 1040 TCCAGGGGGA AACGCCTGGT ATCTTATAG TCCTGTCGGG 1080 TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT 1120 CGTCAGGGGG GCGAGCCTA TCACTCTAGA AACTGCCCT ATAAAAGGGCC AGCCTTAGACC TCACTCTAGA AACTGCCCT ATAAAAGGGCC AGCCTTAGACC TCACTCTAGA AACTGCCCT ATAAAAGGGCC AGCCTTAGACC TCACTCTAGA AACTGCCCT ATAAAAGGGCC AGCCTTAGACC TCACTCTAGA AACTGCCCT ATAAAAGGGCC AGCCTGACC AACCC TCACTCTAGA TCACTCTAGA TCACTCTAGA TCACTCTAGA TCACTCTAGA TCACTCTAGA TCACTCTAGA TCACTCTAGA TCACTCTAGA TC	CCCGAGCGGC	CAAAGGGAGC				
TTTCAAAAGT TCGAAAGAT TAGTTAGGCC CATACCTCGC CATACCTCGC TCTGCTAATC CTGTTACCAG CAGTGGCGAT AAGTCGTGTC CAGTGGCGAT AAGTCGTGTC CGGATAAGGC CGGATAAGGC CGGATAAGGC GGGGTTCGTG CACACAGCCC CACCGGCTG CACACAGCCC CACCGGATC CACCGAACTG CACCGAACTG GCACCGCTC CCGAAGGGAG CCCGCTTC CCGAAGGGAG AAAGGCGGAC CCGAAGGAC CCGAAGGGAG AAAGGCGGAC TTGAGAAAGC CCGAAGGGAG AAAGGCGGAC TCAGGGGGA TCCAGGGGAG AACGCCTGGT TTTTCGCCACC TCTGACTTGA CGCTCGAACTG GCCCGGGGGAGC TCCGGAAGGC CGCGGGGGAGC TCCGGAGGGAG CGCCGGGACCT TCTGACTTGA CCCGCACGCT TTTTCGCCACC TCTGACTTGA CCCGCAGGGAG TCCCGGAGGC TCCGGAGACC TCCGGAGGAGC TCACTCTAGA TTTGGCCACC CCAGCAACGC CCAGCAACGC TCACTCTAGA TGGAAAAACG CCAGCAACGC TCACTCTAGA TGGAGAGACC TCACTCTAGA TGAGAGAGCA TCCGGAGGAC TCACTCTAGA TGAGAGAGCA TCCGGAGGAC TCACTCTAGA TGAGAGAGCA TCCGGTTTCC TCCGTTTTCT TCCGTTTTCT TCCGTTTTCC CCAATTCCC TTATGATACC TTATGACCC TCCCTTAGGG ACCCTCAGAGGC ACCTGAGCT ACCCCTCAGAGGC ACCGCTGAGCC ACCGCTGAGCC ACCGCGGACC ACCGCCGCACC ACCGCCCTC ACCCCCCC ACCACACGC ACCCGCCTAACCC ACCCGCCTAACCC ACCCGCCTAACCC ACCGCCCTAACCC ACCCGCCTAACCC ACCCGCCTAACCC ACCGCCCTC ACCCGCCTAACCC ACCGCCCTC ACCCCCCC ACCACCCCC ACCACCCCCC ACCACCCCC ACCACC	TCGACTTCGA	AGGTTCGAAT				
TAGTTAGGCC ACCACTICAA GAACTCTGTA GCACCGCTA CATACCTCGC TCTGCTAATC CTGTTACCAG TGGCTGCTGC CAGTGGCGAT AAGTCGTGTC TTACCGGGTT GGACTCAAGA 840 CGATAGTTAC CGGATAAGGC GCAGCGGTCG GGCTGAACGG 880 CACCGAACTG CACACAGCCC AGCTTGGAGC GAACGACCTA 920 CACCGAACTG AGATACCTAC AGCGTGAGCA TTGAGAAAGC 960 GCACGCTTC CCGAAGGGAG AAAGGCGGAC AGGTATCCGG 1000 TAAGCGGCAG GGTCGGAACA GGAGAGCGCA CGAGGGAGCT 1040 TCCAGGGGGA AACGCCTGGT ATCTTTATAG TCCTGTCGGG 1080 TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT 1120 CGTCAGGGGG GCGGAGCCTA TGGAAAAACG CCAGCAACGC 1160 GGCCGGGGGGA TCCGGAGAGC TCACTCTAGA TGAGAGAGCA 1200 GTGAGGGGAG GACAGAGACT CGAATTTCCC GAGCTATTTC 1240 AGTTTTCTTT TCCGTTTTGT GCAATTTCAC TTATGATACC 1280 GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG 1320 GATGCCCCTC AACTGGCCT ATAAAGGGCC AGCCTGAGCT 1360 GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC	TTTCAAAAGT	CCGAAAGAAT				. — -
CATACCTCGC TCTGCTATC CTGTTACCAG TGGCTGTC CAGTGGCGAT AAGTCGTGTC TTACCGGGTT GGACTCAAGA CGATAGTTAC CGGATAAGGC GCAGCGGTCG GGCTGAACGG GGGGTTCGTG CACACAGCCC AGCTTGGAGC GAACGACCTA CACCGAACTG AGATACCTAC AGCGTGAGCA TTGAGAAAGC GCCACGCTTC CCGAAGGGAG AAAGGCCGAC AGGTATCCGG TAAGCGGCAG GGTCGGAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGGA AACGCCTGGT ATCTTTATAG TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT CGTCAGGGGG GCGGAGCCTA TGGAAAAACG CCAGCAACGC GGCCGGGGGA TCCCGGAGACC TCACTCTAGA TGAGAGAGCA GTGAGGGAGA GACAGAGACT CGAATTTCCG GAGCTATTTC AGTTTTCTTT TCCGTTTTGT GCAATTTCCC TTATGATACC GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG GAGCCCCTC AACTGGCCCT ATAAAGGGCC AGCCTGAGCC GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC						
CAGTGGCGAT AAGTCGTGTC TTACCGGGTT GGCTGAACGG GGATAGTTAC CGGATAAGGC GCAGCGGTCG GGCTGAACGG GGGGTTCGTG CACACAGCCC AGCTTGGAGC GAACGACCTA CACCGAACTG AGATACCTAC AGCGTGAGCA TTGAGAAAGC GCCACGCTTC CCGAAGGGAG AAAGGCGGAC AGGTATCCGG TAAGCGGCAG GGTCGGAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGGA AACGCCTGGT ATCTTTATAG TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT CGTCAGGGGG GCGGAGCCTA TGGAAAAACG CCAGCAACGC GGCCGGGGGA TCCGGAGAC TCACTCTAGA TGAGAGAGCA GTGAGGGAGA GACAGAGACT CGAATTTCCG GAGCTATTTC GTGATTTCTTT TCCGTTTTGT GCAATTTCAC TTATGATACC GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG GAGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG GAGCGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC			0101111			
GGGTTCGTG CACACAGCCC AGCTTGGAGC GAACGACCTA CACCGAACTG AGATACCTAC AGCGTGAGCA TTGAGAAAGC GCCACGCTTC CCGAAGGAG AAAGGCGGAC AGGTATCCGG TAAGCGGCAG GGTCGGAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGGA AACGCCTGGT ATCTTTATAG TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT CGTCAGGGGG GCGGAGCCTA TGGAAAAACG CCAGCAACGC GGCCGGGGGA TCCGGAGACC TCACTCTAGA TGAGAGAGCA GTGAGGGAGA GACAGAGACT CGAATTTCCG GAGCTATTTC AGTTTTCTTT TCCGTTTTGT GCAATTTCAC TTATGATACC GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG GAGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG GAGCCCCTC AACTGGCCCT ATAAAGGGCC AGCCTGAGCT GAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC						
CACCGAACTG AGATACCTAC AGCTTGGAGC TAGAGAAAGC GCCACGCTTC CCGAAGGGAG AAAGGCGGAC AGGTATCCGG TAAGCGGCAG GGTCGGAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGGA AACGCCTGGT ATCTTTATAG TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT CGTCAGGGGGA TCCGGAGACC TCACTCTAGA TGAGAGACCA GGCCGGGGGA TCCGGAGACC TCACTCTAGA TGAGAGACCA GTGAGGGAGA GACAGAGACT CGAATTTCCG GAGCTATTTC GTGAGGGAGA GACAGAGACT CGAATTTCCC GAGCTATTTC GGCCAATGCT TCGTTGCTAT TTTGGAAACT CCCCTTAGGG GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG GATGCCCCTC AACTGGCCCT ATAAAGGGCC AGCCTGACCT GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC 1320 GCAGAGGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC					•	
GCCACGCTTC CCGAAGGAG AAAGGCGGAC AGGTATCCGG TAAGCGGCAG GGTCGGAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGGA AACGCCTGGT ATCTTTATAG TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT CGTCAGGGGG GCGGAGCCTA TGGAAAAACG CCAGCAACGC GGCCGGGGGA TCCGGAGACC TCACTCTAGA TGAGAGAGCA GTGAGGGAGA GACAGAGACT CGAATTTCCG GAGCTATTTC AGTTTTCTTT TCCGTTTTGT GCAATTTCAC TTATGATACC GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG GAGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG GATGCCCCTC AACTGGCCCT ATAAAGGGCC AGCCTGACCT GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC	GGGGTTCGTG	CACACAGCCC				
TAAGCGCAG GGTCGGAACA GGAGAGCGCA CGAGGGAGCT 1040 TCCAGGGGA AACGCCTGGT ATCTTTATAG TCCTGTCGGG 1080 TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT 1120 CGTCAGGGGG GCGGAGCCTA TGGAAAAACG CCAGCAACGC 1160 GGCCGGGGGA TCCGGAGACC TCACTCTAGA TGAGAGAGCA 1200 GTGAGGGAGA GACAGAGACT CGAATTTCCG GAGCTATTTC 1240 AGTTTTCTTT TCCGTTTTGT GCAATTTCAC TTATGATACC 1280 GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG 1320 GATGCCCCTC AACTGGCCCT ATAAAGGGCC AGCCTGAGCT 1360 GCAGAGAGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC 1400	CACCGAACTG	AGATACCTAC	AGCGTGAGCA	TTGAGAAAGC		
TAAGCGGCAG GGTCGGAACA GGAGAGCCA CGAGGACCA TCCAGGGGGA AACGCCTGGT ATCTTTATAG TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT CGTCAGGGGG GCGAGCCTA TGGAAAAACG CCAGCAACGC GGCCGGGGGA TCCGGAGAGC TCACTCTAGA TGAGAGAGCA GTGAGGGAGA GACAGAGACT CGAATTTCCG GAGCTATTTC AGTTTTCTTT TCCGTTTTGT GCAATTTCAC TTATGATACC GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG GATGCCCCTC AACTGGCCCT ATAAAGGGCC AGCCTGAGCT GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC 1080 108	GCCACGCTTC	CCGAAGGGAG	AAAGGCGGAC	AGGTATCCGG	· · · · · · · · · · · · · · · · · · ·	
TCCAGGGGGA AACGCCIGGI AICITIAIAS ICCIGIGGGGGGGA AACGCCIGGI AICITIAIAS ICCIGIGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	TAAGCGGCAG	GGTCGGAACA	GGAGAGCGCA	CGAGGGAGCT		
CGTCAGGGG GCGAGCCTA TGGAAAAACG CCAGCAACGC 1160 GGCCGGGGGA TCCGGAGAGC TCACTCTAGA TGAGAGAGCA 1200 GTGAGGGAGA GACAGAGACT CGAATTTCCG GAGCTATTCC 1240 AGTTTTCTTT TCCGTTTTGT GCAATTTCAC TTATGATACC 1280 GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG 1320 GATGCCCCTC AACTGGCCCT ATAAAGGGCC AGCCTGAGCT 1360 GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC 1400						
GGCCGGGGA TCCGGAGACT TGGAAAAACG CCAGCAACGC GGCCGGGGA TCCGGAGACC TCACTCTAGA TGAGAGAGCA 1200 GTGAGGGAGA GACAGAGACT CGAATTTCCG GAGCTATTTC 1240 AGTTTTCTTT TCCGTTTTGT GCAATTTCAC TTATGATACC 1280 GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG 1320 GATGCCCCTC AACTGGCCCT ATAAAGGGCC AGCCTGAGCT 1360 GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC 1400						
GTGAGGAGA TCCGGAGAGC TCACTCTAGA TAGACACTAGA TGCGCGGGGAGA TCCGGAGAGC TCACTCTAGA TGCGCCATTTCC TCGATTTCC GAGCTATTCC TATGATACC 1240 GGCCAATGCT TCGTTTGT GCAATTTCAC TTATGATACC 1280 GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG 1320 GATGCCCCTC AACTGGCCCT ATAAAGGGCC AGCCTGAGCT 1360 GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC 1400						
GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG GATGCCCCTC AACTGGCCCT ATAAAGGGCC AGCCTGAGCT GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC 1280 1280						
GGCCAATGCT TGGTTGCTAT TTTGGAAACT CCCCTTAGGG 1320 GATGCCCCTC AACTGGCCCT ATAAAGGGCC AGCCTGAGCT 1360 GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC 1400	GTGAGGGAGA	GACAGAGACT				
GATGCCCTC AACTGGCCCT ATAAAGGGCC AGCCTGAGCT GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC 1360 GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC 1400	AGTTTTCTTT	TCCGTTTTGT	GCAATTTCAC	TTATGATACC		1280
GATGCCCTC AACTGGCCCT ATAAAGGGCC AGCCTGAGCT GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC 1360 GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC 1400						
GCAGAGGATT CCTGCAGAGG ATCAAGACAG CACGTGGACC 1400	GGCCAATGCT	TGGTTGCTAT	TTTGGAAACT	CCCCTTAGGG		
GCAGAGGATT CCIGCAGAGG AICAAGACAG CACGIGGMEC	GATGCCCCTC	AACTGGCCCT	ATAAAGGGCC	AGCCTGAGCT		
TCGCACAGCC TCTCCCACAG GTACC 1425	GCAGAGGATT	CCTGCAGAGG	ATCAAGACAG	CACGTGGACC		
	TCGCACAGCC	TCTCCCACAG	GTACC			1425

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 719 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1911 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

			CCTCCTCCCA	40
TATGAGCAAG	GGCGAGGAAC	TGTTCACTGG	CCCCACAAAT	80
ATTCTCGTGG	AACTGGATGG	CGATGIGAAI	CCACATACCC	120
TTTCTGTCAG	CGGAGAGGGT	GAAGGIGAIG	TCCANACCTC	160
AAAGCTCACC	CTGAAATTCA	TCTGCACCAC	TGGAAAGCIC	200
CCTGTGCCAT	GGCCAACACT	GGTCACTACC	ATATCAACCIAIG	240
GCGTGCAGTG	CTTTTCCAGA	TACCCAGACC	CCCCTATCTC	280
GCATGACTTT	TTCAAGAGCG	CCATGCCCGA	CCCAACTACA	320
CAGGAGAGAA	CCATCTTTTT	CAAAGAIGAC	A C A C C C T C C T	360
AGACCCGCGC	TGAAGTCAAG GAGTTGAAGG	TICGAAGGIG	TARCCCIGGI	400
GAATAGAATC	TCGGCCACAA	GCALIGACII	A A CT A TA A CT	440
GGAAACATTC	GTACATCATG	CCCCACAACC	AAAAAAATGG	480
CCCACAATGT	AACTTCAAGA	TCACACAAGC	CATTCACGAT	520
CATCAAGGTC	AGCTGGCCGA	CCACACACAA	CACIDACOMI	560
GGATCCGTGC	CGGCCCTGTG	CTCCTCCCAG	ACAACCATTA	600
CAATCGGCGA	CAGTCTGCCC	CTCCTCCAG	CCCAACGAAA	640
CCTGTCCACC	CAGTCTGCCC	CTCCACTTC	TGACCGCTGC	680
AGAGAGACCA	CATGGTCCTG	ACCACCTCTA	CAACTGAGCC	720
TGGGATCACA	TGGATGCCTT	CACCCCACCC	CCCCATCCCC	760
ATATGGTGAG	GTCAGAGCCC	CCCCCCAGGC	CAGGCCAATG	800
GAGACCTGTA	CCCTGCAGTG	ACTACTCACT	CCCCCCCTCC	840
CCCGTCCTTC	ACCCCTCCCC	AGTAGTGACT	CTGGCCCTGG	880
GATCCCTGTG	TCCAGTGCCC	ACCACCCTTC	ТССТДАТАДА	920
AAGTTGCCAC	ATCATTTTGT	CTCACTACCT	GTCCTCTATA	960
ATTAAGTTGC	CTTGATATCG	A A TT CTTTTCT	CAACGTAACA	1000
ATATTATAAG	GCGCGTCATT	TCATATCATC	CGCCCCCCCTT	1040
CTTTACAGCG	GAGCAGGCCA	CTANACCAT	TACCCGTGGT	1080
CCCGATAAGG	AGCGGCCAAA	CCCACCACAC	TCTAAATCTG	1120
GGGGTTCCCG	CTTCGAAGGT	TCCA ATCCTT	CCCCCACCAC	1160
CCGTCATCGA	AAAAGTCCGA	AAGAATTCCTT	GCAGCCCGTG	1200
CATCACTITC	TAGGCCACCA	CTTCDDCDDC	TCTGTAGCAC	1240
TAGCCGTAGI	CCTCGCTCTG	CTICAACAAC	TACCAGTGGC	1280
CGCCTACATA	GGCGATAAGT	CCTCTCTTAC	CGGGTTGGAC	1320
TGCTGCCAGI	AGTTACCGGA	TAAGGCGCAG	CGGTCGGGCT	1360
TCAAGACGAI	TTCGTGCACA	CAGCCCAGCT	TGGAGCGAAC	1400
GAACGGGGG	GAACTGAGAT	ACCTACAGCG	TGAGCATTGA	1440
GACCIACACC	CGCTTCCCGA	AGGGAGAAAG	GCGGACAGGT	1480
AMACCACTA	CGGCAGGGTC	GGAACAGGAG	AGCGCACGAG	1520
AICCGGIAAG	GGGGGAAACG	CCTGGTATCT	TTATAGTCCT	1560
CTCCCCTTTC	GCCACCTCTG	ACTTGAGCGT	CGATTTTTGT	1600
CATCCTCCTC	AGGGGGGCGG	AGCCTATGGA	AAAACGCCAG	1640
CANCECECC	GGGGGATCCG	GAGAGCTCAC	TCTAGATGAG	1680
ACACCACTCA	GGGAGAGACA	GAGACTCGAA	TTTCCGGAGC	1720
	TTCTTTTCCG	TTTTGTGCAA	TTTCACTTAT	1760
CATACCCCCC	: AATGCTTGGT	TGCTATTTTG	GAAACTCCCC	1800
THE COORD	CCCCTCAACT	GGCCCTATAA	AGGGCCAGCC	1840
TAGGGGATG	L VCCOTCTTC	CAGAGGATCA	AGACAGCACG	1880
TGAGCTGCAG	: ACAGCCTCTC	CACACACCTAC	C	1911
IGGACCICGC	. MCHGCCICIC	CCACACOIAC	. •	

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 69 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: unknown

- MOLECULE TYPE: peptide (ii)
- SEQUENCE DESCRIPTION: SEQ ID NO:19: (xi)

Pro Asp Leu Ser Tyr Met Pro Ile Trp Lys Phe Pro 10 5 Asp Glu Glu Gly Ala Cys Gln Pro Cys Pro Ile Asn 20 Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys Gly 30 Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser 45 Ile Ile Ser Ala Val Val Gly Ile Leu Leu Val Val 55 Val Leu Gly Val Val Phe Gly Ile Leu 65

- INFORMATION FOR SEQ ID NO:20: (2)
 - SEQUENCE CHARACTERISTICS: (i)
 - LENGTH: 287 amino acids (A)
 - TYPE: amino acid (B)
 - TOPOLOGY: unknown (D)
 - MOLECULE TYPE: protein (ii)
 - SEQUENCE DESCRIPTION: SEQ ID NO:20: (xi)

Pro Ala Pro Gly Ala Gly Gly Met Val His His Arg His Arg Ser Ser Ser Thr Arg Ser Gly Gly Asp 20 Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu Glu Ala 30 Pro Arg Ser Pro Leu Ala Pro Ser Glu Gly Ala Gly 45 40 Ser Asp Val Phe Asp Gly Asp Leu Gly Met Gly Ala 55 50 Ala Lys Gly Leu Ser Leu Pro Thr His Asp Pro Ser 65 Pro Leu Gln Arg Tyr Ser Glu Asp Pro Thr Val Pro 80 Leu Pro Ser Glu Thr Asp Gly Tyr Val Ala Pro Leu 90 85 Thr Cys Ser Pro Gln Pro Glu Tyr Val Asn Gln Pro 105 100 Asp Val Arg Pro Pro Pro Ser Pro Arg Glu Gly Pro 115 110 Leu Pro Ala Ala Arg Pro Ala Gly Ala Thr Leu Glu 125 Arg Pro Lys Thr Leu Ser Pro Gly Lys Asn Gly Val

140

135

Val Lys Asp Val Phe Ala Phe Gly Gly Ala Val Glu 150 Asn Pro Glu Tyr Leu Thr Pro Gln Gly Thr Cys Ser 165 160 Pro Gln Pro Glu Tyr Val Asn Gln Pro Asp Val Arg 175 170 Pro Gln Pro Pro Ser Pro Arg Glu Gly Pro Leu Pro 185 Ala Ala Arg Pro Ala Gly Ala Thr Leu Glu Arg Pro 200 Lys Leu Ser Pro Gly Lys Asn Gly Val Val Lys Asp 215 210 Val Phe Ala Phe Gly Gly Ala Val Glu Asn Pro Glu 220 Tyr Leu Thr Pro Gln Gly Gly Ala Ala Pro Gln Pro 235 His Pro Pro Pro Ala Phe Ser Pro Ala Phe Asp Asn 250 245 Leu Tyr Tyr Trp Asp Asp Pro Pro Glu Arg Gly Ala 255 Pro Pro Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu 270 Asn Pro Glu Tyr Leu Gly Leu Asp Val Pro Val 285 280

- (2) INFORMATION FOR SEQ ID NO:21:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: unknown
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Ile Ile Ser Ala Val Val Gly Ile Leu Leu Val Val 1 5 10

Val Leu Gly Val Val Phe Gly Ile Leu Ile 15 20

- (2) INFORMATION FOR SEQ ID NO:22:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2125 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

- 87 -

(ii) MOLECULE	TYPE:	cDNA
---------------	-------	------

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION:SEQ ID NO:22:

GCCACCATGG	CCCCTGACCT	CTCCTACATG	CCCATCTGGA		40
ACTTTCCAGA	TGAGGAGGGC	GCATGCCAGC	CTTGCCCCAT		80
CAACTGCACC	CACTCCTGTG	TGGACCTGGA	TGACAAGGGC		120
TGCCCCGCCG	AGCAGAGAGC	CAGCCCTCTG	ACGTCCATCA		160
TCTCTGCGGT	GGTTGGCATT	CTGCTGGTCG	TGGTCTTGGG		200
CCTCCTCTTT	GGGATCCTCA	TCAAGCGACG	GCAGCAGAAG		240
ATCACATCTC	CAGACCCTGC	CCCGGGCGCT	GGGGGCATGG		280
TCCACCACAG	GCACCGCAGC	TCATCTACCA	GGAGTGGCGG		320
TCCCCACCTC	ACACTAGGGC	TGGAGCCCTC	TGAAGAGGAG		360
CCCCCCAGGT	CTCCACTGGC	ACCCTCCGAA	GGGGCTGGCT		400
$CCC\Delta TCT\Delta TT$	TGATGGTGAC	CTGGGAATGG	GGGCAGCCAA		440
CCCCCTCCAA	AGCCTCCCCA	CACATGACCC	CAGCCCTCTA		480
CAGCGGTACA	GTGAGGACCC	CACAGTACCC	CTGCCCTCTG		520
ACACTCATCC	CTACGTTGCC	CCCCTGACCT	GCAGCCCCCA		560
CCCTCDDTDT	GTGAACCAGC	CAGATGTTCG	GCCCCAGCCC		600
CCTTCGCCCC	GAGAGGGCCC	TCTGCCTGCT	GCCCGACCTG		640
СТССТСССАС	TCTGGAAAGG	CCCAAGACTC	TCTCCCCAGG		680
CAACAATCCC	GTCGTCAAAG	ACGTTTTTGC	CTTTGGGGGT		720
CCCCTCCAGA	ACCCCGAGAC	TTGACACCCC	AGGGAGGAGC		760
TGCCCCTCAG	CCCCACCCTC	CTCCTGCCTT	CAGCCCAGCC		800
ጥጥሮርልሮልልሮር	TCTATTACTG	GGACCAGGAC	CCACCAGAGC		840
CCCCCCCTCC	ACCCAGCACC	TTCAAAGGGA	CACCTACGGC		880
AGAGAACCCA	GAGTACCTGG	GTCTGGACGT	GCCAGTGTGA		920
AGCCTTAAGG	GCCATATGGT	GAGTGGATGC	CTTGACCCCA		960
GGCGGGGATG	GGGGAGACCT	GTAGTCAGAG	CCCCCGGGCA	_	.000
GCACAGGCCA	ATGCCCGTCC	TTCCCCTGCA	GTGAGTAGTG		.040
A CTCCCCGGG	TGGGATCCCT	GTGACCCCTC	CCCAGTGCCT		.080
CTCCTGGCCC	TGGAAGTTGC	CACTCCAGTG	CCCACCAGCC	_	.120
ጥጥረጥሮሮጥል ልጥ	AAAATTAAGT	TGCATCATTT	TGTCTGACTA		160
CCTCTCTCT	ATAATATTAT	AAGCTTGATA	TCGAATTCTT		200
TCTCDACGTA	ACACTTTACA	GCGGCGCGTC	ATTTGATATG		240
ATCCCCCCC	CTTCCCGATA	AGGGAGCAGG	CCAGTAAAAG	1	L280
CATTACCCCT	GGTGGGGTTC	CCGAGCGGCC	AAAGGGAGCA	-	L320
CALIACCOGI	CTGCCGTCAT	CGACTTCGAA	GGTTCGAATC	1	L360
CTTCCCCCAC	CACCATCACT	TTCAAAAGTC	CGAAAGAATT		L400
CTTCCCCCAC	GTGTAGCCGT	AGTTAGGCCA	CCACTTCAAG	1	L440
A A CTCTCTCTAG	CACCGCCTAC	ATACCTCGCT	CTGCTAATCC	_	L480
TOTTA COACT	GGCTGCTGCC	AGTGGCGATA	AGTCGTGTCT	1	1520
TOT TACCAGE	GACTCAAGAC	GATAGTTACC	GGATAAGGCG	1	1560
CACCCCTCCC	GCTGAACGGG	GGGTTCGTGC	ACACAGCCCA	1	1600
CAGCGGTCGC	AACGACCTAC	ACCGAACTGA	GATACCTACA	1	1640
GCTTGGAGCG	TGAGAAAGCG	CCACGCTTCC	CGAAGGGAGA	1	1680
A ACCCCCA CA	GGTATCCGGT	AAGCGGCAGG	GTCGGAACAG	=	1720
AAGGCGGACA	GAGGGAGCTT	CCAGGGGGAA	ACGCCTGGTA	=	1760
GAGAGCGCAC	CCTGTCGGGT	TTCGCCOCA	CTGACTTGAG	-	1800
TCTTTATAGT	TGTGATGCTC	GTCAGGGGGG	CGGAGCCTAT	:	1840
CGTCGATTTT	CAGCAACGCG	GTCAGGGGGG	CCGGAGAGCT		1880
GGAAAAACGC	GAGCAACGCG GAGAGAGCAG	TCACCCAGACAC	ACAGAGACTC	:	1920
CACTCTAGAT	GAGAGAGCAG GAGCTATTTCA		СССТТТТСТС		1960
GAATTTCCGC	AGCIAITICA	GIIIICIIII	5001111010		

TTGGA	AAACTC EGGCCA EACAGC	TATGATACCG GCCAATGCTT GGTTGCTATT CCCTTAGGGG ATGCCCCTCA ACTGGCCCTA GCCTGAGCTG CAGAGGATTC CTGCAGAGGA ACGTGGACCT CGCACAGCCT CTCCCACAGG	2000 2040 2080 2120 2125
(2)	INFORM	ATION FOR SEQ ID NO:23:	
	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 27 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown	
	(ii)	MOLECULE TYPE: oligonucleotide	
	(iii)	HYPOTHETICAL: No	
	(iv)	ANTI-SENSE: No	
	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:23:	
GTCT	GCCACC	ATGGCCTACT CCCCTGC	27
(2)	INFOR	MATION FOR SEQ ID NO:24:	
	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 36 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	(ii)	MOLECULE TYPE: oligonucleotide	
	(iii)	HYPOTHETICAL: No	
	(iv)	ANTI-SENSE: No	
	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:24:	
TTCI	TTGGTG	ACCTACCTCT TCGGAATTGC CGAGTC	36
(2)	INFOR	MATION FOR SEQ ID NO:25:	
	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 1242 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown	

- 89 -

(ii) MOLECULE	TYPE:	cDNA
---------------	-------	------

- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ATGGAGGAGC	CGCAGTCAGA	TCCTAGCGTC	GAGCCCCCTC		40
TGAGTCAGGA	AACATTTTCA	GACCTATGGA	AACTACTTCC	_	80
TGAAAACAAC	GTTCTGTCCC	CCTTGCCGTC	CCAAGCAATG	-	.20
GATGATTTGA	TGCTGTCCCC	GGACGATATT	GAACAATGGT	_	160
TCACTGAAGA	CCCAGGTCCA	GATGAAGCTC	CCAGAATGCC		200
AGAGGCTGCT	CCCCGCGTGG	CCCCTGCACC	AGCAGCTCCT	-	240
ACACCGGCGG	CCCCTGCACC	AGCCCCCTCC	TGGCCCCTGT	-	085
CATCTTCTGT		AAAACCTACC	AGGGCAGCTA		320
CGGTTTCCGT	CTGGGCTTCT	TGCATTCTGG		-	360
TCTGCCACCA	TGGCCTACTC	CCCTGCGTCT			100
CGTACTCCCC	and the second s	AAGATGTTTT	GCCAACTGGC		140
CAAGACCTGC	CCTGTGCAGC	TGTGGGTTGA	TTCCACACCC		180
CCGCCCGGCA	CCCGCGTCCG	CGCCATGGCC	ATCTACAAGC	_	520
AGTCACAGCA	CATGACGGAG	GTTGTGAGGC	GCTGCCCCCA		560
CCATGAGCGC	TGCTCAGATA	GCGATGGTCT			600
CAGCGTCTTA	TCCGAGTGGA	AGGAAATTTG	CGTGTGGAGT		640
	CAGAAACACT	TTTCGACATA	GTGTGGTGGT		680
GCCCTATGAG	CCGCCTGAGG		CTGTACCACC		720
ATCCACTACA	ACTACATGTG	TAACAGTTCC	TGCATGGGCG		760
GCATGAACCG	GAGGCCCATC	CTCACCATCA	TCACACTGGA		800
AGACTCCAGT	GGTAATCTAC		CAGCTTTGAG		840
GTGCGTGTTT	GTGCCTGTCC		CGGCGCACAG		088
AGGAAGAGAA	TCTCCGCAAG	AAAGGGGAGC	CTCACCACGA		920 960
GCTGCCCCCA	GGGAGCACTA	AGCGAGCACT	GCCCAACAAC		000
ACCAGCTCCT	CTCCCCAGCC	AAAGAAGAAA	CCACTGGATG	-	
GAGAATATTT	CACCCTTCAG	ATCCGTGGGC	GTGAGCGCTT	-	040 080
CGAGATGTTC	TTTGGTGACC	TACCTCTTCG	GAATTGCCGA	_	120
GTCTTCCGAG	AGCTGAATGA	GGCCTTGGAA	CTCAAGGATG		160
CCCAGGCTGG	GAAGGAGCCA	GGGGGGAGCA	GGGCTCACTC	_	200
CAGCCACCTG	AAGTCCAAAA	AGGGTCAGTC	TACCTCCCGC	_	240
CATAAAAAAC	TCATGTTCAA	GACAGAAGGG	CCTGACTCAG	-	240
AC				1	242

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 608 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No

- 90 -

(iv)	ANTI-SENSE:	No
------	-------------	----

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CTCGGGCCGC CCCTGACGAG TGGCGAAACC CCCCTGGAAG GCCGCTTACC AGCGTGGCGC GTTCGGTGTA CGAACCCCCC AACTATCGTC CGCCACTGGC GAGGTATGTA CCTAACTACG GCGCTCTGCT TAGCTCTTGA	CATCACAAAA CGACAGGACT CTCCCTCGTG GGATACCTGT TTTCTCAATG GGTCGTTCGC GTTCAGCCCG TTGAGTCCAA AGCAGCCACT GGCGGTGCTA GCTACACTAG GAAGCCAGTT TCCGGCAAAC	ACCGCTGCGC CCCGGTAAGA GGTAACAGGA CAGAGTTCTT AAGGACAGTA ACCTTCGGAA AAACCACCGC	AAGTCAGAGG CAGGCGTTTC TTCCGACCCT CCCTTCGGGA AGGTATCTCA GCTGTGTGCA CTTATCCGGT CACGACTTAT TTAGCAGAGC GAAGTGGTGG TTTGGTATCT AAAGAGTTGG TGGTAGCGGT	12 16 20 24 28 32 36 40 44 48 52	50 10 30 20 50 40 80 20
	TCCGGCAAAC		TGGTAGCGGT	60	60 00 08

(2) INFORMATION FOR SEQ ID NO:27:

(i)	CECTIENCE	CHARACTERISTICS:

- (A) LENGTH: 1547 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GGTACCTGCC	A CCATCCCCC	GGATTCTTTA	TCACTGATAA	40
	ATATTATGTT	TATCAGTGAT	AAAGTGTCAA	80
01100100110	GTTGCAGCCG	AATACAGTGA	TCCGTGCCGG	120
GCATORICIEE.	TTGAACGAGG	TCGGCGTAGA	CGGTCTGACG	160
CCCTGGACTG	TGGCGGAACG	GTTGGGGGTG	CAGCAGCCGG	200
ACACGCAAAC	GCACTTCAGG	AACAAGCGGG	CGCCTTAAGG	240
CGCTTTACTG	GAGTGGATGC	CTTGACCCCA	GGCGGGGATG	280
GCCATATGGT	GTAGTCAGAG	CCCCCGGGCA	GCACAGGCCA	320
GGGGAGACCT	TTCCCCTGCA		TGAGTGCCTC	360
ATGCCCGTCC	GGAAGTTGCC	ACTCCAGTGC	CCACCAGCCT	400
TCCTGGCCCT	AAATTAAGTT	GCATCATTTT	GTCTGACTAG	440
TGTCCTAATA	TAATATTATA	AGCTTGATAT	CGAATTCTTT	480
GTGTCCTCTA	AAAGTGATGG	TGGTGGGGGA	AGGATTCGAA	520
CGGACTTTTG	CGATGACGGC	AGATTTAGAG	TCTGCTCCCT	560
CCTTCGAAGT		CCACGGGTAA		600
TTGGCCGCTC	GGGAACCCCA		ATCATATCAA	640
GGCCTGCTCC	CTTATCGGGA	GTTACGTTGA		680
ATGACGCGCC	GCTGTAAAGT	GGCGTTTTTC		720
CTGCAGCCCG	CCGCGTTGCT	GGCGIIIIC	CHIMOCIEC	

GCCCCCTGA	CGAGCATCAC	AAAAATCGAC	GCTCAAGTCA	760
	AACCCGACAG	GACTATAAAG	ATACCAGGCG	800
TTTCCCCCTG	GAAGCTCCCT	CGTGCGCTCT	CCTGTTCCGA	840
CCCTGCCGCT	TACCGGATAC	CTGTCCGCCT	TTCTCCCTTC	880
GGGAAGCGTG	GCGCTTTCTC	AATGCTCACG	CTGTAGGTAT	920
CTCAGTTCGG	TGTAGGTCGT	TCGCTCCAAG	CTGGGCTGTG	960
TGCACGAACC	CCCCGTTCAG	CCCGACCGCT	GCGCCTTATC	1000
CGGTAACTAT	CGTCTTGAGT	CCAACCCGGT	AAGACACGAC	1040
TTATCGCCAC	TGGCAGCAGC	CACTGGTAAC	AGGATTAGCA	1080
GAGCGAGGTA	TGTAGGCGGT	GCTACAGAGT	TCTTGAAGTG	1120
GTGGCCTAAC	TACGGCTACA	CTAGAAGGAC	AGTATTTGGT	1160
ATCTGCGCTC	TGCTGAAGCC	AGTTACCTTC	GGAAAAAGAG	1200
TTGGTAGCTC	TTGATCCGGC	AAACAAACCA	CCGCTGGTAG	1240
CGGTGGTTTT	TTTGTTTGCA	AGCAGCAGAT	TACGCGCAGA	1280
AAAAAAGGAT	CTGGGGGATC	CGGAGAGCTC	ACTCTAGATG	1320
AGAGAGCAGT	GAGGGAGAGA	CAGAGACTCG	AATTTCCGGA	1360
GCTATTTCAG	TTTTCTTTTC	CGTTTTGTGC	AATTTCACTT	1400
ATGATACCGG	CCAATGCTTG	GTTGCTATTT	TGGAAACTCC	1440
CCTTAGGGGA	TGCCCCTCAA	CTGGCCCTAT	AAAGGGCCAG	1480
CCTGAGCTGC	AGAGGATTCC	TGCAGAGGAT	CAAGACAGCA	1520
CGTGGACCTC	GCACAGCCTC	TCCCACA		1547
			•	

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1807 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GGTACCTGCC	ACCATGGCGC	GGATTCTTTA	TCACTGATAA	4	ŧΟ
		TATCAGTGAT		8	30
	GTTGCAGCCG	AATACAGTGA		12	30
CCCTGGACTG	TTGAACGAGG	TCGGCGTAGA	CGGTCTGACG	16	50
	TGGCGGAACG		CAGCAGCCGG	20	0 (
CGCTTTACTG		AACAAGCGGG	CGCCTTAAGG	24	ŧΟ
GCCATATGGT	GAGTGGATGC			28	3 O
• • • • • • • •	GTAGTCAGAG		GCACAGGCCA	32	3 O
		GGATGAGTAG	TGAGTGCCTC	36	50
		ACTCCAGTGC		40	0(
		GCATCATTTT		4.4	ŧΟ
GTGTCCTCTA	TAATATTATA	AGCTTGATAT	CGAATTCTTT	4.8	3 O
CGGACTTTTG	AAAGTGATGG	TGGTGGGGGA	AGGATTCGAA	52	30
CCTTCGAAGT	CGATGACGGC	AGATTTAGAG	TCTGCTCCCT	56	50
TTGGCCGCTC		CCACGGGTAA	TGCTTTTACT	60	0 (
GGCCTGCTCC	CTTATCGGGA	AGCGGGGCGC	ATCATATCAA	64	
ATGACGCGCC	GCTGTAAAGT	GTTACGTTGA	GAAAGAATTC	68	30

CTGCAGCCCG	CCGCGTTGCT	GGCGTTTTTC	CATAGGCTCC	•	20
GCCCCCTGA	CGAGCATCAC	AAAAATCGAC	GCTCAAGTCA		60
GAGGTGGCGA	AACCCGACAG	GACTATAAAG	ATACCAGGCG	•	00
TTTCCCCCTG	GAAGCTCCCT	CGTGCGCTCT	CCTGTTCCGA	_	40
CCCTGCCGCT	TACCGGATAC		TTCTCCCTTC	•	80
GGGAAGCGTG	GCGCTTTCTC	AATGCTCACG	CTGTAGGTAT	-	20
CTCAGTTCGG	TGTAGGTCGT	TCGCTCCAAG	CTGGGCTGTG	_	60 100
TGCACGAACC	CCCCGTTCAG	CCCGACCGCT	GCGCCTTATC		40
CGGTAACTAT	CGTCTTGAGT	CCAACCCGGT	AAGACACGAC		80
TTATCGCCAC	TGGCAGCAGC	CACTGGTAAC	AGGATTAGCA		20
GAGCGAGGTA		GCTACAGAGT	TCTTGAAGTG		.60
GTGGCCTAAC		CTAGAAGGAC	AGTATTTGGT	- -	200
ATCTGCGCTC		AGTTACCTTC			240
TTGGTAGCTC		AAACAAACCA	CCGCTGGTAG		280
CGGTGGTTTT		AGCAGCAGAT	TACGCGCAGA		320
AAAAAAGGAT	CTGGGGGATC	CGGAGAGCTC	CCAACGCGTT		360
					100
ATAAGCAGGA	ACTTTGAAGA	CTCAGTGACT	CAGTGAGTAA		140
TAAAGACTCA	GTGACTTCTG	ATCCTGTCCT	AACTGCCACT		180
CCTTGTTGTC		GGCTTCCTGC		 -	520
GGACCCCTTC	CCTGGAAGGT	AAAACTAAGG			560
AGAAATTTTT	CCACCATTGG		AAGAGGAAAC GAGGGAGAGA		500
TGATGAGCTC					540
CAGAGACTCG	AATTTCCGGA				580
CGTTTTGTGC			• • • • • • • • • • • • • • • • • • • •		720
GTTGCTATTT				_	760
CTGGCCCTAT	AAAGGGCCAG	CCTGAGCTGC		-	800
TGCAGAGGAT	CAAGACAGCA	CGTGGACCTC	GCACAGCCTC	- :	807
TCCCACA					

INFORMATION FOR SEQ ID NO:29: (2)

- SEQUENCE CHARACTERISTICS: (i)
 - LENGTH: 2308 base pairs TYPE: nucleic acid (A)
 - (B)
 - STRANDEDNESS: single (C)
 - TOPOLOGY: unknown (D)
- MOLECULE TYPE: CDNA (ii)
- HYPOTHETICAL: No (iii)
- ANTI-SENSE: No (iv)
- SEQUENCE DESCRIPTION: SEQ ID NO: 29: (xi)

GGTACCTGCC ACCATGGCGA	AGGGCGAGGA	ACTGTTCACT	40
GGCGTGGTCC CAATTCTCGT	CCAACTGGAT	GGCGATGTGA	80
ATGGGCACAA ATTTTCTGTC	AGCGGAGAGG	GTGAAGGTGA	120
TGCCACATAC GGAAAGCTCA	CCCTGAAATT	CATCTGCACC	160
ACTGGAAAGC TCCCTGTGCC	ATGGCCAACA	CTGGTCACTA	200
CCTTCACCTA TGGCGTGCAG	TGCTTTTCCA	GATACCCAGA	240
CCTTCACCTA TGGCGTGCAC	TTTTCAAGAG	CGCCATGCCC	280

GAGGGCTATG	TGCAGGAGAG	AACCATCTTT	TTCAAAGATG		320
ACGGGAACTA	CAAGACCCGC	GCTGAAGTCA	AGTTCGAAGG		360
TGACACCCTG	GTGAATAGAA	TCGAGTTGAA	GGGCATTGAC		400
TTTAAGGAAG	ATGGAAACAT	TCTCGGCCAC	AAGCTGGAAT		440
ACAACTATAA	CTCCCACAAT	GTGTACATCA	TGGCCGACAA		480
			GATCAGACAC		520
	ATGGATCCGT		GACCATTATC		560
AACAGAACAC	TCCAATCGGC	GACGGCCCTG	TGCTCCTCCC		600
AGACAACCAT	TACCTGTCCA	CCCAGTCTGC	CCTGTCTAAA		640
			CTGCTGGAGT		680
			TGGACGAGCT		720
	GCGCCTTAAG		TGAGTGGATG		760
CCTTGACCCC	AGGCGGGGAT	GGGGGAGACC	TGTAGTCAGA		800
GCCCCCGGGC	AGCACAGGCC	AATGCCCGTC	CTTCCCCTGC		840
AGGATGAGTA	GTGAGTGCCT	CTCCTGGCCC	TGGAAGTTGC		880
	CCCACCAGCC		AAAATTAAGT		920
	TGTCTGACTA		ATAATATTAT		960
			GAAAGTGATG		1000
			TCGATGACGG		1040
CAGATTTAGA	GTCTGCTCCC	TTTGGCCGCT	CGGGAACCCC		1080
ACCACGGGTA	ATGCTTTTAC	TGGCCTGCTC	CCTTATCGGG		1120
AAGCGGGGCG	CATCATATCA	AATGACGCGC	CGCTGTAAAG	•	1160
	AGAAAGAATT		GCCGCGTTGC		1200
TGGCGTTTTT	CCATAGGCTC	CGCCCCCTG	ACGAGCATCA		1240
CAAAAATCGA	CGCTCAAGTC	AGAGGTGGCG	AAACCCGACA		1280
	GATACCAGGC		GGAAGCTCCC		1320
			TTACCGGATA		1360
CCTGTCCGCC	TTTCTCCCTT	CGGGAAGCGT	GGCGCTTTCT		1400
	GCTGTAGGTA		GTGTAGGTCG		1440
	GCTGGGCTGT		CCCCCGTTCA		1480
GCCCGACCGC	TGCGCCTTAT	CCGGTAACTA	TCGTCTTGAG		1520
	TAAGACACGA		CTGGCAGCAG		1560
CCACTGGTAA	CAGGATTAGC	AGAGCGAGGT	ATGTAGGCGG		1600
	TTCTTGAAGT		CTACGGCTAC		1640
ACTAGAAGGA	CAGTATTTGG	TATCTGCGCT	CTGCTGAAGC		1680
CAGTTACCTT	CGGAAAAAGA	GTTGGTAGCT	CTTGATCCGG		1720
CAAACAAACC	ACCGCTGGTA	GCGGTGGTTT	TTTTGTTTGC		1760
	TTACGCGCAG		TCTGGGGGAT		1800
	CCCAACGCGT		GATGAGGGAA		1840
	ATCTGTAATG		AACTTTGAAG		1880
ACTCAGTGAC	TCAGTGAGTA	ATAAAGACTC	AGTGACTTCT		1920
GATCCTGTCC	TAACTGCCAC	TCCTTGTTGT	CCCAAGAAAG		1960
	CTCTCTGAGG		CCCTGGAAGG		2000
	GATGTCAGCA		TCCACCATTG		2040
GTGCTTGGTC	AAAGAGGAAA	CTGATGAGCT	CACTCTAGAT		2080
GAGAGAGCAĢ	TGAGGGAGAG	ACAGAGACTC	GAATTTCCGG		2120
AGCTATTTCA	GTTTTCTTTT	CCGTTTTGTG	CAATTTCACT		2160
	GCCAATGCTT				2200
CCCTTAGGGG	ATGCCCCTCA	ACTGGCCCTA	TAAAGGGCCA		2240
GCCTGAGCTG	CAGAGGATTC	CTGCAGAGGA	TCAAGACAGC		2280
ACGTGGACCT	CGCACAGCCT	CTCCCACA			2308